Hong Kong University of Science and Technology

COMP170: Discrete Mathematical Tools for Computer Science

Spring 2009

Final Exam

21 May 2009, 12:30–3:00pm, Rm 3007

Solutions

Question 1: (a) True.

We prove the identity algebraically. Since

$$n^{\underline{k}} = \frac{n!}{(n-k)!}$$
$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

we obtain the identity $n^{\underline{k}} = \binom{n}{k}k!$.

(b) False.

The fact that $f : A \to B$ is injective means that

$$\forall x, y \in A \Big(x \neq y \Rightarrow f(x) \neq f(y) \Big).$$

So B should have at least as many elements as A, i.e., $|B| \geq |A|$.

(c) False.

The two numbers p and q in the RSA algorithm must be prime, but 3534 is not a prime number.

(d) False.

The quantified statement is true as long as there exists a non-CS student x who makes CS(x) false, even though there does not exist any Computer Science student who gets an A+ grade. The correct statement should be:

$$\exists x \in U \Big(\mathrm{CS}(x) \wedge \mathrm{APlus}(x) \Big).$$

(e) True.

By definition,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

So,

$$\begin{aligned}
P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\
&= \frac{P(B|A)P(A)}{P(B \cap (A \cup \overline{A}))} \\
&= \frac{P(B|A)P(A)}{P((B \cap A) \cup (B \cap \overline{A}))} \\
&= \frac{P(B|A)P(A)}{P(B \cap A) + P(B \cap \overline{A})} \qquad \text{(because } B \cap A \text{ and } B \cap \overline{A} \text{ are disjoint)} \\
&= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})}.
\end{aligned}$$

Question 2: (a) The problem can be re-stated as follows:

> Let $m = 20$ and $n = 9$ be two positive integers such that $\gcd(m, n) = \gcd(20, 9) = 1$ and we have the following two modular equations:
>
> $$x \bmod 20 = 13 \tag{1}$$
> $$x \bmod 9 = 0. \tag{2}$$
>
> Find the value of x in the range $100 \le x \le 200$.

From the Chinese Remainder Theorem, we know that there exists a unique solution for x in $Z_{mn} = Z_{180}$. Let us first find this solution and then check if it satisfies the condition $100 \le x \le 200$.

Let $m = 20, n = 9, a = 13, b = 0$, $\overline{m}$ is the multiplicative inverse of m in $Z_n = Z_9$, and $\overline{n}$ is the multiplicative inverse of n in $Z_m = Z_{20}$.

From the constructive proof for the Chinese Remainder Theorem, the unique solution for x in $Z_{mn} = Z_{180}$ can be computed as

$$x = y \bmod mn = y \bmod 180$$

where

$$y = an\overline{n} + bm\overline{m} = 13 \cdot 9 \cdot \overline{n} = 117 \cdot \overline{n}.$$

Since

$$9 \cdot 9 \bmod 20 = 1,$$

we have $\overline{n} = 9$. Thus

$$y = 117 \cdot 9 = 1053$$

and

$$x = 1053 \bmod 180 = 153.$$

Because $100 \le 153 \le 200$, so $x = 153$ is the solution.

(b) We note that

$$x = 153 + kmn = 153 + 180k$$

for any integer k must satisfy both (1) and (2).

Since the only value of k that can make x satisfy the condition $500 \le x \le 600$ is when $k = 2$, so the value of x is

$$x = 153 + 360 = 513.$$

Question 3: (a) We first use an n-tuple $(G_1, \ldots, G_i, \ldots, G_n)$ to denote a pairing, where each element $G_i \subset \{1, \ldots, 2n\}$, with $|G_i| = 2$, contains two numbers that correspond to two of the $2n$ participants who play together in a game. The number of n-tuples that correspond to different pairings is

$$\binom{2n}{2}\binom{2n-2}{2}\binom{2n-4}{2}\cdots\binom{2}{2} = \frac{(2n)!}{2^n}.$$

However, note that the order of the elements in an n-tuple is important but it is not important when considering a pairing for the games. Because each pairing is counted $n!$ times using an (ordered) n-tuple representation, the total number of ways to arrange the pairing should be

$$\frac{(2n)!}{2^n n!}.$$

(b) Since

$$P(n) = \frac{(2n)!}{2^n n!}$$
$$P(n-1) = \frac{(2n-2)!}{2^{n-1}(n-1)!},$$

we obtain

$$P(n) = \frac{2n(2n-1)}{2n} P(n-1) = (2n-1)P(n-1).$$

Question 4: (a) We rewrite $S(1), \ldots, S(5)$ as follows:

$$\begin{aligned} S(1) &= 2 - 1 \\ S(2) &= 2 \cdot 3 - 1 \\ S(3) &= 2 \cdot 3 \cdot 4 - 1 \\ S(4) &= 2 \cdot 3 \cdot 4 \cdot 5 - 1 \\ S(5) &= 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 - 1. \end{aligned}$$

From the above we guess that

$$S(n) = (n+1)! - 1.$$

(b) We prove the correctness of the closed form

$$S(n) = (n+1)! - 1$$

by induction.

- **Base case:** $(n = 1)$
  Since $1 \cdot 1! = 1 = (1+1)! - 1$, the closed form holds for $n = 1$.
- **Inductive step:** $(n > 1)$
  Assuming that the closed form holds for $n - 1$ as the inductive hypothesis, we want to show that it also holds for n.
  By applying the inductive hypothesis,

$$\begin{aligned} S(n) &= \sum_{k=1}^{n} k \cdot k! \\ &= S(n-1) + n \cdot n! \\ &= n! - 1 + n \cdot n! \\ &= (n+1) \cdot n! - 1 \\ &= (n+1)! - 1. \end{aligned}$$

  So the closed form also holds for n.

By the principle of mathematical induction, we show that the closed form holds for all positive integers n.

Question 5: (a) By iterating the recurrence, we get

$$\begin{aligned}
T(n) &= aT(n/m) + bn^d \\
&= a\left[aT(n/m^2) + \frac{bn^d}{m^d}\right] + bn^d \\
&= a^2T(n/m^2) + \frac{abn^d}{m^d} + bn^d \\
&= a^2\left[aT(n/m^3) + \frac{bn^d}{m^{2d}}\right] + \frac{abn^d}{m^d} + bn^d \\
&= a^3T(n/m^3) + \frac{a^2bn^d}{m^{2d}} + \frac{abn^d}{m^d} + bn^d \\
&= a^3T(n/m^3) + bn^d \sum_{i=0}^{2} \left(\frac{a}{m^d}\right)^i \\
&\quad \vdots \\
&= a^{\log_m n}T(1) + bn^d \sum_{i=0}^{(\log_m n)-1} \left(\frac{a}{m^d}\right)^i \\
&= cn^{\log_m a} + bn^d \cdot \frac{1-(a/m^d)^{\log_m n}}{1-a/m^d} \qquad \text{(since } a \neq m^d\text{)} \\
&= cn^{\log_m a} + bn^d \cdot \frac{1-\frac{n^{\log_m a}}{n^d}}{1-a/m^d} \\
&= cn^{\log_m a} + \frac{bm^d}{m^d-a}\left(n^d - n^{\log_m a}\right) \\
&= \frac{cm^d - bm^d - ac}{m^d-a} n^{\log_m a} + \frac{bm^d}{m^d-a} n^d.
\end{aligned}$$

(b) We want to prove the correctness of the following formula, where $n = m^i$, for all integers $i \geq 0$:

$$T(n) = \frac{cm^d - bm^d - ac}{m^d-a} n^{\log_m a} + \frac{bm^d}{m^d-a} n^d.$$

- **Base case:** $(i = 0)$
  Because

$$\begin{aligned}
T(n) = T(m^0) = T(1) &= \frac{cm^d - bm^d - ac}{m^d-a} \cdot 1^{\log_m a} + \frac{bm^d}{m^d-a} \cdot 1^d \\
&= \frac{cm^d - bm^d - ac}{m^d-a} + \frac{bm^d}{m^d-a} \\
&= c,
\end{aligned}$$

  the formula holds for the base case.

- **Inductive step:** $(i > 0)$
  As the inductive hypothesis, we assume that the formula for $T(n)$ holds for $n = m^{i-1}$. We want to show that it also holds for $n = m^i$.

By applying the inductive hypothesis to the recurrence, we obtain

$$\begin{aligned}
T(n) &= aT(n/m) + bn^d \\
&= a\left[\frac{cm^d - bm^d - ac}{m^d - a}\left(\frac{n}{m}\right)^{\log_m a} + \frac{bm^d}{m^d - a}\left(\frac{n}{m}\right)^d\right] + bn^d \\
&= a\left[\frac{cm^d - bm^d - ac}{m^d - a} \cdot \frac{n^{\log_m a}}{m^{\log_m a}} + \frac{bm^d}{m^d - a} \cdot \frac{n^d}{m^d}\right] + bn^d \\
&= \frac{cm^d - bm^d - ac}{m^d - a} n^{\log_m a} + \frac{abn^d}{m^d - a} + bn^d \\
&= \frac{cm^d - bm^d - ac}{m^d - a} n^{\log_m a} + \frac{bm^d}{m^d - a} n^d.
\end{aligned}$$

Thus $T(n)$ holds for $n = m^i$.

By the principle of mathematical induction, we show that the formula for $T(n)$ holds for all n which are nonnegative integer powers of m.

(c) By substituting

$$m = 3,\ a = 2,\ b = 1,\ c = 1,\ d = 2$$

into the closed formula obtained in part (a), we get

$$\begin{aligned}
T(n) &= \frac{3^2 - 3^2 - 2}{3^2 - 2} n^{\log_3 2} + \frac{3^2}{3^2 - 2} n^2 \\
&= \frac{9}{7} n^2 - \frac{2}{7} n^{\log_3 2}.
\end{aligned}$$

Question 6: Our first attempt is to rewrite the statement as follows before devising an inductive proof:

> If $T(n) \leq 9T(n/3) + n$, then there exist n_0 and $k > 0$ such that $T(n) \leq kn^2$ for all $n = 3^i \geq n_0$ where i is a nonnegative integer.

After some experimentation, we find that the inductive hypothesis is not strong enough to let us prove what we want. So we strengthen the inductive hypothesis by rewriting the statement to prove as follows:

> If $T(n) \leq 9T(n/3) + n$, then there exist n_0 and $k_1, k_2 > 0$ such that $T(n) \leq k_1n^2 - k_2n$ for all $n = 3^i \geq n_0$ where i is a nonnegative integer.

We first check if choosing $n_0 = 3^0 = 1$ allows us to yield a proof. For the base case to hold, we want

$$T(1) \leq k_1 - k_2. \tag{3}$$

Now, we assume inductively that there exist k_1, k_2 such that

$$T(m) \leq k_1m^2 - k_2m$$

for $m < n$. Then

$$\begin{aligned} T(n) &\leq 9T(n/3) + n \\ &\leq 9\left[k_1\left(\frac{n}{3}\right)^2 - k_2\left(\frac{n}{3}\right)\right] + n \\ &= k_1n^2 - 3k_2n + n \\ &= (k_1n^2 - k_2n) - 2k_2n + n. \end{aligned}$$

If we choose k_2 such that $-2k_2n + n \leq 0$ or $k_2 \geq \frac{1}{2}$, then we have $T(n) \leq k_1n^2 - k_2n$.

Let us choose $k_2 = 1$. From (3), we have $k_1 \geq T(1) + k_2 = T(1) + 1$. With these values for k_1, k_2 and $n_0 = 1$, we have completed an inductive proof for the statement.

Therefore, $T(n) = O(n^2)$.

Alternatively, you may also solve this by iterating the recurrence inequality.

Question 7: (a) If only one of the four tossers says YES and the other three say NO, there will be five YES votes and six NO votes. Since the last member will go along with the majority (NO), the defendant will not be found guilty.

On the other hand, if two tossers say YES and two say NO, there will be six YES votes and five NO votes. Since the last member will go along with the majority (YES), the defendant will be found guilty.

Obviously, if more than two tossers say YES, the defendant will also be found guilty.

So at least two tossers are needed to say YES for the defendant to be found guilty.

(b) Because each of the four tossers tosses a fair coin, we have a total of 2^4 possible outcomes each of the same probability, i.e., uniform distribution.

Of the 2^4 possible outcomes, there are $\binom{4}{k}$ $(0 \leq k \leq 4)$ ways for k of them to say YES.

So the probability that the defendant is found guilty is equal to

$$1 - \frac{\binom{4}{0} + \binom{4}{1}}{2^4} = 1 - \frac{1+4}{2^4} = \frac{11}{16}.$$

Question 8: (a) The events are not pairwise disjoint because the intersection of any two distinct events is not empty.

(b) Since the probability distribution defined on S is uniform, we can express $P(E_i)$ as

$$P(E_i) = \frac{|E_i|}{|S|}.$$

(c) Applying the inclusion-exclusion principle, we get

$$\begin{aligned} P\left(\bigcup_{i=1}^{100} E_i\right) &= \sum_{1\le i\le 100} P(E_i) - \sum_{1\le i<j\le 100} P(E_i \cap E_j) + \sum_{1\le i<j<k\le 100} P(E_i \cap E_j \cap E_k) \\ &= \binom{100}{1}\frac{1000}{|S|} - \binom{100}{2}\frac{200}{|S|} + \binom{100}{3}\frac{6}{|S|} \\ &= \frac{100000 - 990000 + 970200}{|S|} \\ &= \frac{80200}{|S|} \end{aligned}$$

Since $P(\bigcup_{i=1}^{100} E_i) = P(S) = 1$, we have $|S| = 80200$ $(< 100000 = \sum_{1\le i\le 100} |E_i|)$. So

$$P(E_i) = \frac{|E_i|}{|S|} = \frac{1000}{80200} = \frac{5}{401}.$$

Question 9: For each of the 100 questions, let the ordered pair (P, S) denote the correct answer P assigned by the professor and the answer S written down by the student, where $P, S \in \{\mathtt{T}, \mathtt{F}\}$.

Since the 100 questions correspond to a Bernoulli trials process, the expected values and variances can be computed as follows:

$$\begin{aligned}
E(S_1) &= 100 \cdot P(\{(\mathtt{T},\mathtt{T}), (\mathtt{F},\mathtt{F})\}) \\
&= 100 \cdot \Big[P(\{(\mathtt{T},\mathtt{T})\}) + P(\{(\mathtt{F},\mathtt{F})\})\Big] \\
&= 100 \cdot \left(\frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{1}{2}\right) \\
&= 100 \cdot \frac{1}{2} \\
&= 50 \\
E(S_2) &= 100 \cdot \Big[P(\{(\mathtt{T},\mathtt{T})\}) + P(\{(\mathtt{F},\mathtt{F})\})\Big] \\
&= 100 \cdot \left(0 + \frac{2}{3} \cdot 1\right) \\
&= 100 \cdot \frac{2}{3} \\
&= \frac{200}{3} \\
V(S_1) &= 100 \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) \\
&= 25 \\
V(S_2) &= 100 \cdot \frac{2}{3} \cdot \left(1 - \frac{2}{3}\right) \\
&= \frac{200}{9}.
\end{aligned}$$

Question 10: (a) For the random variable X_i, $i-1$ locations of the hash table have been filled and we want to count the number of items to hash before one more empty location is filled. For each trial, the probability of success is

$$p_i = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}.$$

Thus,

$$E(X_i) = \frac{1}{p_i} = \frac{n}{n-i+1}.$$

(b) By the linearity of expectation, we get

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} E(X_i) \\ &= \sum_{i=1}^{n} \frac{n}{n-i+1} \\ &= n \sum_{i=1}^{n} \frac{1}{n-i+1} \\ &= n \sum_{i=1}^{n} \frac{1}{i} \\ &= nH_n, \end{aligned}$$

where H_n is the nth harmonic number. (Note that this is just the *coupon collector's problem.*)

(c)

$$\begin{aligned} V(Y) &= E((Y-E(Y))^2) \\ &= E(Y^2 - 2YE(Y) + E(Y)^2) \\ &= E(Y^2) - E(Y)^2. \end{aligned}$$

(d) Since the probability that $X_i = k$ (and hence $X_i^2 = k^2$), for $k \in \{1, 2, \ldots\}$, is equal to $(1-p_i)^{k-1} p_i$ ($k-1$ failures followed by a success), the expected value $E(X_i^2)$ can be computed as follows:

$$\begin{aligned} E(X_i^2) &= \sum_{k=1}^{\infty} (1-p_i)^{k-1} p_i \, k^2 \\ &= \frac{p_i}{1-p_i} \sum_{k=1}^{\infty} (1-p_i)^k \, k^2 \\ &= \frac{p_i}{1-p_i} \cdot \frac{(1-p_i)(2-p_i)}{p_i^3} \qquad \text{(identity 4 on page 2 of exam)} \\ &= \frac{2-p_i}{p_i^2}. \end{aligned}$$

From part (c), we can compute $V(X_i)$ as follows:

$$\begin{aligned} V(X_i) &= E(X_i^2) - E(X_i)^2 \\ &= \frac{2-p_i}{p_i^2} - \frac{1}{p_i^2} \\ &= \frac{1-p_i}{p_i^2} \\ &= \frac{n(i-1)}{(n-i+1)^2}. \end{aligned}$$

(e) Because the random variables $X_i, 1 \leq i \leq n$, are independent, we can compute $V(X)$ as follows:

$$\begin{aligned} V(X) &= \sum_{i=1}^{n} V(X_i) \\ &= n \sum_{i=1}^{n} \frac{i-1}{(n-i+1)^2} \\ &= n \sum_{i=1}^{n-1} \frac{n-i}{i^2}. \end{aligned}$$

Question 11: Let $P_1, P_2, \ldots, P_n$ denote the n points.

If all n points lie along a semicircular arc, then there must exist a point, say P_i, such that the semicircular arc starting at P_i and going clockwise around the circle contains no other point $P_j, j \neq i$. Let E_i denote such an event. The probability of E_i is

$$P(E_i) = \frac{1}{2^{n-1}},$$

because each of the $n-1$ points other than P_i can only lie on half of the circumference.

We note that if there exists a point P_i that satisfies the event E_i, then there does not exist a different point P_j that satisfies the corresponding event E_j. This implies that the n events are disjoint.

Hence, the desired probability is

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) = \frac{n}{2^{n-1}}.$$